

statistics

MATHEMATISCHE OPERATIONSFORSCHUNG UND STATISTIK

SERIES STATISTICS

Volume 8

1977

Number 1-4

Published at

Central Institute of Mathematics and Mechanics
Academy of Sciences of the G.D.R.

Editorial Board

January/December

J. Anděl (Prague), H. Bandemer (Freiberg),
R. Bartoszyński (Warsaw), H. Bunke (Berlin),
O. Bunke (Berlin), V. Dupač (Prague),
V. V. Fedorov (Moscow), B. W. Gnedenko (Moscow),
J. Jurečková (Prague), W. Klonecki (Wrocław),
E. Lyttkens (Uppsala), G. M. Mania (Tbilissi),
P.-H. Müller (Dresden), B. I. Penkov (Sofia),
M. L. Puri (Bloomington, Indiana), C. R. Rao (New Delhi),
P. Révész (Budapest), J. Roy (Calcutta),
K. Sarkadi (Budapest), J.-L. Soler (Grenoble),
E. Spjøtvoll (Aas), T. Văduva (Bukarest),
I. Vincze (Budapest), W. Winkler (Dresden),
H. Wold (Göteborg)

Chief Editor:

OLAF BUNKE
assisted by J. Gladitz

AKADEMIE-VERLAG · BERLIN



On Hsu's Model in Regression Analysis

FRIEDRICH PUKELSHEIM¹

Summary. The paper exemplifies with Hsu's model a general pattern as how to derive results of variance component estimation from well known results of mean estimation, as far as linear model theory is concerned. This '*dispersion-mean-correspondence*' provides new and short proofs for various theorems from the literature, concerning unbiased invariant quadratic estimators with minimum BAYES risk or minimum variance. For pure variance component models, unbiased non-negative quadratic estimability is characterized in terms of the design matrices.

1. Introduction

The purpose of this communication is to exemplify with Hsu's model a general pattern as how to derive results of variance component estimation from known results of mean estimation.

The dispersion-mean-correspondence (Sect. 2) introduces a derived model such that mean regression in the derived model corresponds to estimating the variance components in the original model. Hsu's model specifies the fourth moments via the kurtosis γ , thus opening the way for minimum variance estimation. Sect. 3 collects some matrix algebra for convenient reference. In Sect. 4, some known results of J. KLEFFE and R. PINCUS [8], P. L. HSU [6], H. DRYGAS [2], and C. R. RAO [14] are proved by persistently applying the dispersion-mean-correspondence. This approach extends insight and understanding of linear model theory, providing short proofs, slight generalizations, and alternative characterizations. Sect. 5 is concerned with the existence of unbiased non-negative definite quadratic estimates of variance components and presents an estimability criterion in terms of the design matrices that specify the model.

Previous Work. S. K. MITRA [10] suggests an approach that is very close to the dispersion-mean-correspondence as presented here or in [12], he, however, stops exploitation at an intermediate stage. J. SEELY [16] and other authors, cf., H. DRYGAS [3], S. GNOT, W. KLONECKI and R. ZMYSLONY [4], J. KLEFFE [7], R. ZMYSLONY [18], reduce variance component estimation to mean estimation in coordinate free terms which seems to provoke ready application somewhat less than the dispersion-mean-correspondence. Most of the examples in Sect. 4 were originally proved by explicitly minimizing the risk function, though the connection to mean estimation is hinted at (H. DRYGAS [2, p. 382]) or

¹ Institut für Mathematische Stochastik der Universität, Hermann-Herder-Str. 10, D-7800 Freiburg im Breisgau, Federal Republic of Germany.

implicit (C. R. RAO [14, p. 451]). The algebraic notions as introduced below have successfully been utilized by other authors as well, cf., J. KLEFFE and R. PINCUS [8], S. K. MITRA [10], G. P. H. STYAN [17], R. ZMYSLONY [18].

Notations. For a matrix A , let A' , A^+ , $\Re A$, $\Re^1 A$ denote its transposed matrix, MOORE-PENROSE inverse [15, p. 26], range (column space), and the range's orthogonal complement, respectively. Let $\text{vec } A$ be the vector obtained from A by ordering its entries lexicographically. The KRONECKER product [15, p. 29] is denoted by \otimes . The function vec is an inner product and tensor product preserving vector space isomorphism:

$$(\text{vec } A)' \text{vec } B = \text{trace } AB', \quad (1)$$

$$\text{vec } xy' = x \otimes y, \quad (2)$$

as follows by considering the standard basis vectors e_v with v -th component 1 and zeroes elsewhere, and the basis matrices $E_{v\mu} = e_v e_\mu'$ with (v, μ) -entry 1 and zeroes elsewhere. For $(\text{vec } A)'$ we shall also write $\text{vec}' A$. I denotes an identity matrix, its order following from the context.

2. Model Set-up and the Dispersion-Mean-Correspondence

The General Linear Model. For an \mathbb{R}^n -valued random vector Y , a linear model is specified by linear decompositions of both the mean vector EY and the dispersion matrix (variance covariance matrix) DY :

$$EY = \sum_{\pi=1}^p b_\pi x_\pi = Xb, \quad DY = \sum_{\kappa=1}^k t_\kappa V_\kappa, \quad M = I - XX^+, \quad (3)$$

where the (n, p) -matrix $X = [x_1 : \dots : x_p]$ and the k symmetric (n, n) -matrices V_κ are known, whereas $b = (b_1, \dots, b_p)'$ and $t = (t_1, \dots, t_k)'$ are to be estimated; M is the orthogonal projector onto $\Re^1 X$.

For estimating t or linear functions of t , we choose, as usual, quadratic estimators $Q(Y)$ which, by definition, are derived from bilinear functions $B(.,.)$ by setting both arguments equal to Y : $Q(Y) = B(Y, Y)$. A maximal invariant statistic with respect to all 'mean translations' $y \rightarrow y + Xb$, $b \in \mathbb{R}^p$, is MY (cf., J. SEELY [16, p. 1646], J. KLEFFE [7]). Thus, $Q(Y)$ is an *invariant quadratic estimator* (IQE) iff $Q(Y) = Q(MY)$.

E.g., when estimating a linear form $q't$, $q \in \mathbb{R}^k$, the set of all IQEs is $\{Y'AY\}$ where A is an arbitrary symmetric (n, n) -matrix satisfying $A = MAM$, or, equivalently, $AX = 0$.

Invariance is a natural statistical requirement: All of Y which is in the range $\Re X$ may be explained by mean regression, leaving the residuals MY for inference on the dispersion parameter t . Technically speaking, invariant estimates of t are free from the mean parameter, cf., R. R. CORBEIL and S. R. SEARLE [1], J. KLEFFE [7]. Finally, the expectation of a quadratic estimate $Y'AY$ does not depend on the mean parameter iff $X'AX = 0$. In most applications, as in HSU's model, A should be non-negative definite (NND). This, however, and $X'AX = 0$ imply $AX = 0$, i.e., invariance.

The Dispersion-Mean-Correspondence. For a linear model (3), consider the derived random R^{n^2} -vector $\mathbf{MY} \otimes \mathbf{MY}$. By (3) and (2), $\mathbf{EMY} \otimes \mathbf{MY} = \mathbf{M} \otimes \mathbf{M} \cdot \text{vec} \sum t_{\alpha} \mathbf{V}_{\alpha}$, and $\mathbf{MY} \otimes \mathbf{MY}$ gives rise to a linear model for mean estimation. Introducing the (n^2, k) -matrices

$$\mathbf{D} = [\text{vec } \mathbf{V}_1 : \cdots : \text{vec } \mathbf{V}_k], \quad \mathbf{D}_M = \mathbf{M} \otimes \mathbf{M} \cdot \mathbf{D}, \quad (4)$$

we arrive at

$$\mathbf{EMY} \otimes \mathbf{MY} = \mathbf{D}_M \mathbf{t}. \quad (5)$$

Thus, \mathbf{t} may be looked at as the dispersion parameter in the original model (3), or as the mean parameter in the derived model (5); this we call the *dispersion-mean-correspondence*.

It remains to be shown that the class of natural estimators of \mathbf{t} is not changed by the dispersion mean-correspondence. Clearly, for any (k, n^2) -matrix \mathbf{L} ,

$$\mathbf{Q}(\mathbf{Y}) = \mathbf{L} \cdot \mathbf{MY} \otimes \mathbf{MY} \quad (6)$$

as a linear estimator of \mathbf{t} in the derived model (5), is an IQE of \mathbf{t} in the original model (3). Conversely, the **KRONECKER** product is a tensor product [5, p. 12], i. e., for every bilinear function $B(\mathbf{x}, \mathbf{y})$ there exists a (unique) linear function \mathbf{L} such that $B(\mathbf{x}, \mathbf{y}) = \mathbf{L} \cdot \mathbf{x} \otimes \mathbf{y}$. This implies that any IQE $\mathbf{Q}(\mathbf{Y})$ is representable as in (6), and thus is a linear estimator in the derived model. E.g., for an IQE $\mathbf{Y}'\mathbf{A}\mathbf{Y}$ of a linear form $\mathbf{q}'\mathbf{t}$ we get from (1) and (2) $\mathbf{Y}'\mathbf{A}\mathbf{Y} = \text{trace } \mathbf{A} \cdot \mathbf{MY}(\mathbf{MY})' = \text{vec}' \mathbf{A} \cdot \mathbf{MY} \otimes \mathbf{MY}$, i.e., $\mathbf{L} = \text{vec}' \mathbf{A}$.

In Sect. 4 we shall be concerned with minimum variance estimation of \mathbf{t} . To this end, we introduce the (n^2, n^2) -matrices

$$\mathbf{F} = \mathbf{D}\mathbf{Y} \otimes \mathbf{Y}, \quad \mathbf{F}_M = \mathbf{M} \otimes \mathbf{M} \cdot \mathbf{F} \cdot \mathbf{M} \otimes \mathbf{M}, \quad \mathbf{N} = \mathbf{I} - \mathbf{D}_M \mathbf{D}_M^+, \quad (7)$$

i.e., the matrix of all central mixed fourth moments of \mathbf{Y} , the dispersion matrix of $\mathbf{MY} \otimes \mathbf{MY}$, and the orthogonal projector onto $\Re^+ \mathbf{D}_M$, respectively. The following lemma is stated for later reference. It applies the celebrated **LEHMANN-SCHEFFÉ** theorem [15, p. 317] to the derived model; the alternative representations in part (ii) follow by (1), (4), and (12).

Lemma 2.1. *Let a linear model be given by (3) and (7); let \mathbf{L} be a (k, n^2) -matrix, \mathbf{A} be a symmetric (n, n) -matrix, and $\mathbf{q} \in \mathbb{R}^k$. Then:*

- (i) $\mathbf{L} \cdot \mathbf{MY} \otimes \mathbf{MY}$ is an unbiased IQE of \mathbf{t} with minimum variance under \mathbf{F}_M (among all other unbiased IQE) iff $\mathbf{L}\mathbf{D}_M = \mathbf{I}$ and $\mathbf{L}\mathbf{F}_M\mathbf{N} = \mathbf{0}$.
- (ii) $\mathbf{Y}'\mathbf{A}\mathbf{Y}$ is an unbiased IQE of $\mathbf{q}'\mathbf{t}$ with minimum variance under \mathbf{F}_M iff $\mathbf{A}\mathbf{X} = \mathbf{0}$, $\mathbf{q} = \mathbf{D}_M' \cdot \text{vec } \mathbf{A} = (\text{trace } \mathbf{V}_1 \mathbf{A}, \dots, \text{trace } \mathbf{V}_k \mathbf{A})'$, and

$$\mathbf{F}_M \cdot \text{vec } \mathbf{A} \in \Re \mathbf{D}_M = \{ \text{vec } \sum \lambda_{\alpha} \mathbf{M} \mathbf{V}_{\alpha} \mathbf{M} / \lambda \in \mathbb{R}^k \}.$$

Hsu's Model. Hsu's model specifies the fourth moments \mathbf{F} via (quasi-)independence and the kurtosis γ of k random effects ξ_{α} , cf., H. DRYGAS [2], P. L. HSU [6], J. KLEFFE and R. PINCUS [8], C. R. RAO [14]. Following C. R. RAO [14, p. 446] we assume a linear decomposition according to

$$\mathbf{Y} - \mathbf{E}\mathbf{Y} = \sum_{\alpha=1}^k \mathbf{U}_{\alpha} \xi_{\alpha} = \mathbf{U} \xi, \quad (8)$$

where the \mathbf{U}_{α} are known (n, c_{α}) -matrices, $c = \sum c_{\alpha}$, $\mathbf{U} = [\mathbf{U}_1 : \cdots : \mathbf{U}_k]$ is of order (n, c) , and $\xi = [\xi_1' : \cdots : \xi_k']'$ is a random \mathbb{R}^c -vector whose independent subvectors

ξ_κ have independent components $\xi_{\kappa,v}$ satisfying

$$E\xi_{\kappa,v}=0, \quad E\xi_{\kappa,v}^2=\sigma_\kappa^2, \quad E\xi_{\kappa,v}^4=(\gamma_\kappa+3)\sigma_\kappa^4, \quad v=1, \dots, c_\kappa. \quad (9)$$

In general, a random \mathbb{R}^c -vector ξ whose independent components satisfy $E\xi_v=0$, $E\xi_v^2=\bar{\sigma}_v^2$, $E\xi_v^4=(\bar{\gamma}_v+3)\bar{\sigma}_v^4$, has mixed fourth moments

$$D\xi \otimes \xi = \sum_{\mu, \nu=1}^c \bar{\sigma}_\mu^2 \bar{\sigma}_\nu^2 (E_{\mu\mu} \otimes E_{\nu\nu} + E_{\mu\nu} \otimes E_{\nu\mu} + \delta_{\mu\nu} \bar{\gamma}_\nu E_{\nu\nu} \otimes E_{\nu\nu}). \quad (10)$$

Finally, we introduce

$$H = \begin{bmatrix} 1 & & & \\ c'_1 & & & 0 \\ & \ddots & & \\ 0 & & 1 & \\ & & c'_k & \end{bmatrix}, \quad \bar{\sigma}^2 = H' \sigma^2, \quad \bar{\gamma} = H' \gamma, \quad (11)$$

where the vectors $\mathbf{1}_{c_\kappa}$ consist of c_κ ones, so that H is of order (k, c) . Thus, (10) and (11) yield the mixed fourth moments under assumption (9). We are now ready to precisely define a HSU-model:

Definition. A HSU-model is a linear model as specified by (3), (4), (7), (10), (11), where $V_\kappa = U_\kappa U'_\kappa$, $\kappa=1, \dots, k$; U and the U_κ 's are as in (8), $t = \sigma^2 = (\sigma_1^2, \dots, \sigma_k^2)'$, $\gamma = (\gamma_1, \dots, \gamma_k)'$, and $F_M = F_M(\sigma^2, \gamma) = MU \otimes MU \cdot (D\xi \otimes \xi) \cdot U'M \otimes U'M$. In order to appealingly display $F_M(\sigma^2, \gamma)$, we now collect some matrix algebra.

3. Some Matrix Algebra

The Separating Property of *vec*. For any 3 matrices A, B, C of appropriate order one has

$$\text{vec } ABC = A \otimes C' \cdot \text{vec } B. \quad (12)$$

Taking $B = E_{\nu\nu}$, this follows at once from (2).

Diagonalizer. For a square matrix A , $\text{Diag } A$ denotes the diagonal matrix with diagonal entries copied from A . For a vector γ , $\text{Diag } \gamma$ is the diagonal matrix with diagonal equal to γ .

Introducing the (c^2, c) -matrix $D_c = [\text{vec } E_{11} : \dots : \text{vec } E_{cc}]$, where E_{11} etc. are basis (c, c) -matrices, yields, by (2), $D_c D'_c = \sum \text{vec } E_{\nu\nu} \cdot \text{vec}' E_{\nu\nu} = \sum (e_\nu \otimes e_\nu) \times \times (e_\nu \otimes e_\nu)' = \sum E_{\nu\nu} \otimes E_{\nu\nu}$, and, by (12), $D_c D'_c \cdot \text{vec } A = \text{vec } \text{Diag } A$. More general, for every diagonal (c, c) -matrix Δ , and for every (c, c) -matrix A , we have

$$D_c \Delta D'_c \cdot \text{vec } A = \text{vec } \Delta \text{Diag } A, \quad \text{where } D_c = [\text{vec } E_{11} : \dots : \text{vec } E_{cc}]. \quad (13)$$

HADAMARD'S Produkt. When diagonalizer are used, HADAMARD's product $A * B = ((A_{ij} \cdot B_{ij}))$ is not far. This is due to $\text{Diag } A = I * A$ and the following lemma.

Lemma 3.1. Let D_c, D_n be defined as in (13) of order (c^2, c) , (n^2, n) , respectively. Let A, B two (c, n) -matrices, and $a \in \mathbb{R}^c, b \in \mathbb{R}^n$. Then:

- (i) $A * B = D'_c \cdot A \otimes B \cdot D_n$.
- (ii) $a' \cdot A * B \cdot b = \text{trace } \text{Diag } a \cdot A \cdot \text{Diag } b \cdot B'$.

Proof. Verification is immediate when taking A, B, a, b to be basis matrices (vectors). ■

Using Lemma 3.1, many properties of the HADAMARD product, cf., e.g., G. P. H. STYAN [17], may easily be inferred from corresponding properties of the KRONECKER product.

4. On Estimates in the Hsu-Model

The Mixed Fourth Moment. In the next two lemmas, we study the mixed fourth moments as a linear operator.

Lemma 4.1. *Let ξ be a random R^c -vector whose independent components ξ_v satisfy $E\xi_v = 0$, $E\xi_v^2 = \bar{\sigma}_v^2$, and $E\xi_v^4 = (\bar{\gamma}_v + 3) \bar{\sigma}_v^4$. For a fixed $\bar{\sigma}_0^2$, put $\Delta_1 = D\xi = \text{Diag } \bar{\sigma}_0^2$, $\Gamma = \text{Diag } \bar{\gamma}$, $\Delta_2 = \Delta_1$, $\Gamma\Delta_1$, and let D_c be defined as in (13). Then, for every (n, n) -matrix A ,*

$$(D\xi \otimes \xi) \cdot \text{vec } A = (2\Delta_1 \otimes \Delta_1 + D_c \Delta_2 D'_c) \cdot \text{vec } \frac{A + A'}{2}. \quad (14)$$

Proof. The assertion is a consequence of formulae (10), (12), and (13). ■

Lemma 4.2. *Assume a Hsu-model. For a fixed σ_0^2 , put $V_* = \sum \sigma_{0,*}^2 V_* = U\Delta_1 U'$. Then:*

(i) *For every (n, n) -matrix A ,*

$$F_M(\sigma_0^2, \gamma) \cdot \text{vec } A = (2MV_*M \otimes MV_*M + MU \otimes MU \cdot D_c \Delta_2 D'_c \cdot U'M \otimes U'M) \cdot \text{vec } \frac{A + A'}{2}. \quad (15)$$

(ii) *For every symmetric (n, n) -matrix A satisfying $A = MAM$,*

$$F_M(\sigma_0^2, \gamma) \cdot \text{vec } A = \text{vec } (2MV_*A V_*M + MU\Delta_2 \cdot \text{Diag } U'AU \cdot U'M). \quad (16)$$

Proof. Part (i) follows from Lemma 4.1, and, by (12), implies part (ii). ■

With these preparations we are now ready to characterize optimal estimates in the Hsu-model.

BAYES estimates. J. KLEFFE and R. PINCUS [8, Th. 3.8] consider unbiased IQEs with minimum BAYES risk:

Theorem 4.1. *Assume a Hsu-model with a priori distribution P for σ^2 , put*

$$\bar{R} = E_P \sigma^2 \sigma'^2, \quad \bar{S} = \text{Diag } H' \bar{R} H, \quad (17)$$

*and let $q \in R^k$. Then $Y'AY$ is an unbiased IQE of $q't$ with minimum BAYES risk at γ iff $AX = O$, $q = (\text{trace } V_1 A, \dots, \text{trace } V_k A)'$, and $MU \cdot (2H' \bar{R} H + \bar{S} \Gamma) * (U'AU) \cdot U'M$ is a linear combination of $MV_1 M, \dots, MV_k M$.*

Proof. By Lemma 4.2.ii, the risk operator is $\Phi(A) = \text{vec } E_P (2MU\Delta_1 U'AU\Delta_1 \cdot U'M + MU\Delta_1^2 \Gamma \cdot \text{Diag } U'AU \cdot U'M)$. But, cf., [17], $\Delta_1 = \text{Diag } H'\sigma^2$ implies $\Delta_1 U'AU\Delta_1 = (H'\sigma^2 \sigma'^2 H) * (U'AU)$ and $\Delta_1^2 = (H'\sigma^2 \sigma'^2 H) * I = \text{Diag } H'\sigma^2 \sigma'^2 H$. Thus $E_P \Delta_1^2 = \bar{S}$, and we get $\Phi(A) = \text{vec } MU \cdot (2H' \bar{R} H + \bar{S} \Gamma) * (U'AU) \cdot U'M$. The assertion now follows by Lemma 2.1.ii, mutatis mutandis. ■

MV estimates. The rest of this section is concerned with *minimum variance unbiased IQE* (MV-UB-IQE) of a linear form $q't$.

P. L. HSU was the first in this area, and his problem [6, Th. 2] reads in terms of the derived model: When is the simple least squares estimate of minimum variance? HSU assumes independent components of γ with equal variances σ^2 and possibly unequal kurtosis $\gamma_1, \dots, \gamma_n$.

Theorem 4.2. Assume a HSU-model with $c=n$, $U=I$, $V_*= \sigma^2 I$, and $s=\text{rank } X < n$. Put $M_2=M * M$, and $m=(M_{11}, \dots, M_{nn})'$. Then $(n-s)^{-1}Y'MY$ is a MV-UB-IQE of σ^2 iff

$$M_2 \Gamma m = \varrho m, \quad \varrho = (n-s)^{-1} m' \Gamma m. \quad (18)$$

Proof. Here, $D_M = \text{vec } M$, so that the simple least squares estimate is $(D'_M D_M)^{-1} D'_M \cdot MY \otimes MY = (\text{trace } M)^{-1} \text{trace } M \cdot MY (MY)' = (n-s)^{-1} Y'MY$, by (1) and (2). This is of minimum variance [11, p. 148] iff $\Re D_M$ is invariant under $F_M(\sigma_0^2, \gamma)$. Lemma 4.2.ii yields $F_M(\sigma_0^2, \gamma) \cdot \text{vec } M = \sigma_0^4 \text{vec } (2M + M\Gamma \cdot \text{Diag } M \cdot M)$. Thus, a necessary and sufficient condition is

$$M \cdot \text{Diag } \Gamma m \cdot M = \varrho M, \quad \varrho = (n-s)^{-1} m' \Gamma m. \quad (19)$$

By considering the diagonal elements, (19) implies (18). The converse follows, with Lemma 3.1.ii, from $\|M \cdot \text{Diag } \Gamma m \cdot M - \varrho M\|^2 = \text{trace } \text{Diag } \Gamma m \cdot M \cdot \text{Diag } \Gamma m \cdot M - 2\varrho \text{trace } \text{Diag } m \cdot \Gamma \cdot \text{Diag } m + \varrho^2 \text{trace } M = m' \Gamma M_2 \Gamma m - 2\varrho m' \Gamma m + \varrho m' \Gamma m = m' \Gamma (M_2 \Gamma m - \varrho m)$. ■

The next result, due to H. DRYGAS [2, Th. 3.5.a], does also lead to eq. (19).

Theorem 4.3. Let $Y'AY$ be an IQE of σ^2 for the HSU-model in Th. 4.2. Then $Y'AY$ is a MV-UB-IQE of σ^2 iff $\text{trace } A=1$ and $2A + M\Gamma \cdot \text{Diag } A \cdot M$ is a scalar multiple of M .

Proof. Using Lemmas 2.1.ii, 4.2.ii, we get $1 = D'_M \text{vec } A = \text{trace } A$, and $F_M(\sigma_0^2, \gamma) \text{vec } A = \sigma_0^4 \text{vec } (2MAM + M\Gamma \cdot \text{Diag } A \cdot M)$. ■

For the general HSU-model, C. R. RAO [14, Th. 1] derives the following estimates. J. KLEFFE [7, Th. 2] gives similar representations.

Theorem 4.4. Let a HSU-model be given. For a fixed σ_0^2 , introduce the (n, n) -matrices $V_* = \sum \sigma_{0,x}^2 V_x = U \Delta_1 U'$, and $R_* = (MV_* M)^+$, and the (c, c) -matrices $M_1 = U' R_* U$, $M_2 = M_1 * M_1$. Let $q \in \mathbb{R}^k$. If V_* is positive definite, and the \mathbb{R}^c -vector ϑ and the \mathbb{R}^k -vector λ satisfy

$$H\vartheta = q, \quad 2\vartheta + M_2 \Delta_2 \vartheta = M_2 H' \lambda, \quad (20)$$

then an unbiased IQE of $q't$ with minimum variance at (σ_0^2, γ) is given by

$$Y' R_* U \cdot \text{Diag } \frac{1}{2} (H' \lambda - \Delta_2 \vartheta) \cdot U' R_* Y. \quad (21)$$

Proof. The assertion follows from Lemma 2.1.ii. Firstly, since $R_* = R_* M$, we have $A = \frac{1}{2} R_* U \cdot \text{Diag } (H' \lambda - \Delta_2 \vartheta) \cdot U' R_*$. Now, $\text{vec } U \cdot \text{Diag } H' \lambda \cdot U' = \text{vec } \sum \lambda_x V_x = D\lambda$, and, by (13), $\text{vec } \Delta_2 \text{Diag } \vartheta = D_c \Delta_2 \vartheta$. Thus, from (12), $L' = \text{vec } A = \frac{1}{2} R_* \otimes R_* \cdot D\lambda - \frac{1}{2} R_* U \otimes R_* U \cdot D_c \Delta_2 \vartheta$. Using $D = U \otimes U \cdot D_c H'$ and Lemma 3.1, this yields $D'_M L' = \frac{1}{2} H D'_c \cdot U' \otimes U' \cdot R_* \otimes R_* \cdot U \otimes U \cdot D_c H' \lambda - \frac{1}{2} H D'_c \cdot M_1 \otimes$

$\otimes M_1 \cdot D_c \Delta_2 \mathfrak{D} = \frac{1}{2} HM_2 H' \lambda - \frac{1}{2} HM_2 \Delta_2 \mathfrak{D}$, which equals q under the assumption (20).

Secondly, the positive definiteness of V_* implies $R_*^+ R_* = M$. Lemma 4.2.i, then, yields $F_M(\sigma_0^2, \gamma) \cdot L' = M \otimes M \cdot D \lambda - MU \otimes MU \cdot D_c \Delta_2 \mathfrak{D} + \frac{1}{2} MU \otimes MU \cdot D_c \Delta_2 D'_c \cdot U' R_* \otimes U' R_* \cdot D \lambda - \frac{1}{2} MU \otimes MU \cdot D_c \Delta_2 D'_c \cdot U' R_* U \otimes U' R_* U \cdot D_c \Delta_2 \mathfrak{D} = D_M \lambda - \frac{1}{2} MU \otimes MU \cdot D_c \Delta_2 (2\mathfrak{D} - M_2 H' \lambda + M_2 \Delta_2 \mathfrak{D})$. Under the assumption (20), this is in $\mathfrak{R}D_M$. ■

Uniformity Criteria. Essentially, the last theorem of this section is also due to C. R. RAO [14, p. 453–454]. The first part characterizes those situations when the estimates are independent of the kurtosis γ ; the second part assumes quasi normality, i.e., $\gamma = 0$, and investigates independence from σ_0^2 .

Theorem 4.5. Assume the Hsu-model and notation of Th. 4.4. Then:

- (i) All MV–UB–IQEs at $(\sigma_0^2, 0)$ are of minimum variance at (σ_0^2, γ) iff $\mathfrak{R}D_M$ is invariant under $MU \otimes MU \cdot D_c \Delta_1^2 \cdot \text{Diag } H' \gamma \cdot D'_c \cdot U' R_* \otimes U' R_*$, or, equivalently, iff $\mathfrak{R}M_2 H'$ is invariant under $M_2 \Delta_1^2 \cdot \text{Diag } H' \gamma$.
- (ii) All MV–UB–IQEs at $(1_k, 0)$ are of minimum variance at $(\sigma_0^2, 0)$ iff $\mathfrak{R}D_M$ is invariant under $MV_* R \otimes MV_* R$, or, equivalently, iff for every $\lambda \in \mathbb{R}^k$ there exist a $\mu \in \mathbb{R}^k$ such that $\sum \lambda_x RV_* RV_* R = \sum \mu_x RV_* R$. Here, $V = \sum V_x$ is assumed positive definite, $R = (MVM)^+$, and V_* is assumed non-negative definite.

Proof. (i) We have to check [11, p. 147] when $\mathfrak{R}F_M(\sigma_0^2, \gamma) \cdot N \subset \mathfrak{R}F_M(\sigma_0^2, 0) \cdot N$. This is the case, by Lemma 4.2.i and after premultiplying with $R_* \otimes R_*$, iff the range of $A = R_* U \otimes R_* U \cdot D_c \Delta_2 D'_c \cdot U' M \otimes U' M \cdot N$ is contained in $\mathfrak{R}M \otimes M \cdot N$. But $M \otimes M \cdot N = N \cdot M \otimes M$ is a projector, so $N \cdot M \otimes M \cdot A = A$ yields $D'_M \cdot R_* U \otimes R_* U \cdot D_c \Delta_2 D'_c \cdot U' M \otimes U' M \cdot N = 0$. This is true iff

$$\mathfrak{R}MU \otimes MU \cdot D'_c \Delta_2 D'_c \cdot U' R_* \otimes U' R_* \cdot D_M \subset \mathfrak{R}D_M, \quad (22)$$

which is the first characterization. Using $D = U \otimes U \cdot D_c H'$ and premultiplying (22) with $D_c \cdot U' R_* \otimes U' R_*$ yields the second characterization, [14, eq. 5.8].

For part (ii), check $\mathfrak{R}F_M(\sigma_0^2, 0) \cdot N \subset \mathfrak{R}F_M(1_k, 0)$ analogously to get the first characterization. Then premultiply with $R \otimes R$, use $D_M \lambda = \text{vec } \sum \lambda_x MV_* M$ and (12) to get the second characterization [14, eq. 6.4]. ■

As indicated in the proof of part (ii), formulae (12) and (13) may be used to reformulate the above criteria in matrix space. The formulation above parallels that of mean estimation: The BLUE under V is BLUE under V_* iff $\mathfrak{R}X$ is invariant under $V_* V^{-1}$ (cf., [11, p. 149]).

For further results on uniform MV–UB–IQE see H. DRYGAS [3] and S. GNOT, W. KLONECKI and R. ZMYSLONY [4]. The dispersion-mean-correspondence may also be used to get MINQUE, weighted least squares, or Ridge-type estimates of variance components, cf., F. PUKELSHEIM [12], [13]. Maximum likelihood estimates are considered, e.g., by R. R. CORBEIL and S. R. SEARLE [1].

5. Unbiased NND Quadratic Estimability

Though one should not dispense with requiring that a quadratic estimate $\mathbf{Y}'\mathbf{A}\mathbf{Y}$ of a single variance component σ_x^2 be non-negative definite (NND), one does, and there are only few investigations of this subject, cf., H. DRYGAS [2], L. R. LA MOTTE [9]. The next lemma is implicitly given by L. R. LA MOTTE [9, p. 728], the formulation below is to stress that NND estimates shift the problem from linearity into convexity.

Lemma 5.1. *Assume a linear model as given by (3), let $\mathbf{q} \in \mathbb{R}^k$. Then:*

(i) *There exists an unbiased IQE of $\mathbf{q}'\mathbf{t}$ iff*

$$\mathbf{q} \in \text{linear hull } \{\mathbf{D}'_{\mathbf{M}} \cdot \mathbf{y} \otimes \mathbf{y} \mid \mathbf{y} \in \mathbb{R}^n\}.$$

(ii) *There exists an unbiased NND quadratic estimator of $\mathbf{q}'\mathbf{t}$ iff*

$$\mathbf{q} \in \text{convex hull } \{\mathbf{D}'_{\mathbf{M}} \cdot \mathbf{y} \otimes \mathbf{y} \mid \mathbf{y} \in \mathbb{R}^n\}.$$

Proof. Unbiasedness and non-negative definiteness imply invariance, see above Sect. 2. $\mathbf{Y}'\mathbf{A}\mathbf{Y}$ is unbiased for $\mathbf{q}'\mathbf{t}$ iff $\mathbf{q} = \mathbf{D}'_{\mathbf{M}} \cdot \text{vec } \mathbf{A}$. Assertions i, ii then follow from the spectral representation [15, p. 39] of symmetric matrices, NND matrices, respectively. ■

What has the model to look like such that a single component t_x be unbiasedly NND estimable? For a pure variance components model, e.g., a Hsu-model, we finally prove as a necessary and sufficient condition: The x -th dispersion design must properly contribute to the explanation of the error space $\Re \mathbf{M} \sum \mathbf{V}_x \mathbf{M}$, (cf., [15, p. 297]).

Theorem 5.1. *For a linear model (3), let all \mathbf{V}_x , $x = 1, \dots, k$, be NND, and fix x . Then there exists an unbiased NND quadratic estimator of t_x iff*

$$\Re \mathbf{M} \mathbf{V}_x \mathbf{M} \not\subset \Re \mathbf{M} \sum_{\lambda \neq x} \mathbf{V}_\lambda \mathbf{M}.$$

Proof. Put $\Re = \text{convex hull } \{\mathbf{D}'_{\mathbf{M}} \cdot \mathbf{y} \otimes \mathbf{y} \mid \mathbf{y} \in \mathbb{R}^n\}$. Clearly, $\Re = \{\sum \mathbf{D}'_{\mathbf{M}} \cdot \mathbf{y}_v \otimes \mathbf{y}_v \mid \mathbf{y}_1, \dots, \mathbf{y}_n \in \mathbb{R}^n\}$. With all \mathbf{V}_x NND, it is easily shown that $\mathbf{e}_x \in \Re$ iff $\mathbf{e}_x \in \{\mathbf{D}'_{\mathbf{M}} \cdot \mathbf{y} \otimes \mathbf{y} \mid \mathbf{y} \in \mathbb{R}^n\}$. The latter means that the nullspace of $\mathbf{M} \sum_{\lambda \neq x} \mathbf{V}_\lambda \mathbf{M}$ be not contained in the nullspace of $\mathbf{M} \mathbf{V}_x \mathbf{M}$. This is the orthogonal dual of the assertion. ■

Acknowledgement

The author would like to thank Prof. H. WITTING for his encouragement and permanent support. He is also grateful to Prof. H. DRYGAS and the referee for suggesting improvements and extensions.

References

- [1] CORBELL, R. R., and S. R. SEARLE (1976). Restricted maximum likelihood (REML) estimation of variance components in the mixed model. *Technometrics* 18, 31–38.
- [2] DRYGAS, H. (1972). The estimation of residual variance in regression analysis. *Math. Operationsforsch. Statist.* 3, 373–388.

- [3] DRYGAS, H. (1977). Best quadratic unbiased estimation in variance-covariance component models. *Math. Operationsforsch. Statist. Ser. Statistics* 8, 211–231.
- [4] GNOT, S., W. KLONECKI and R. ZMYSLONY (1977). Uniformly minimum variance unbiased estimation in various classes of estimators, I. *Math. Operationsforsch. Statist. Ser. Statistics* 8, 199–210.
- [5] GREUB, W. H. (1967). *Multilinear Algebra*. Springer-Verlag, Berlin.
- [6] HSU, P. L. (1938). On the best quadratic estimate of variance. *Statist. Res. Mem.* 2, 91–104.
- [7] KLEFFE, J. (1977). Invariant methods for estimating variance components in mixed linear models. *Math. Operationsforsch. Statist. Ser. Statistics* 8, 233–250.
- [8] KLEFFE, J., and R. PINCUS (1974). BAYES and best quadratic unbiased estimators for variance components and heteroscedastic variances in linear models. *Math. Operationsforsch. Statist.* 5, 147–159.
- [9] LA MOTTE, L. R. (1973). On non-negative quadratic unbiased estimation of variance components. *J. Amer. Statist. Assoc.* 68, 728–730.
- [10] MITRA, S. K. (1971). Another look at RAO's MINQUE of variance components. *Bull. Inst. Internat. Statist.* 44, 2, 279–283.
- [11] MITRA, S. K., and B. J. MOORE (1973). GAUSS-MARKOV estimation with an incorrect dispersion matrix. *Sankhyā Ser. A* 35, 139–152.
- [12] PUKELSHEIM, F. (1977). Estimating variance components in linear models. *J. Multivariate Anal.* 6, 626–629.
- [13] PUKELSHEIM, F. (1977). Equality of two BLUEs and Ridge-type estimates. *Comm. Statist.-Theory Methods A* 6, 603–610.
- [14] RAO, C. R. (1971). Minimum variance quadratic unbiased estimation of variance components. *J. Multivariate Anal.* 1, 445–456.
- [15] RAO, C. R. (1973). *Linear Statistical Inference and its Applications*, Second Edition. John Wiley & Sons, New York.
- [16] SEELY, J. (1972). Completeness for a family of multivariate normal distribution. *Ann. Math. Statist.* 43, 1644–1647.
- [17] STYAN, G. P. H. (1973). HADAMARD products and multivariate statistical analysis. *Linear Algebra and Appl.* 6, 217–240.
- [18] ZMYSLONY, R. (1973). Estimation of variance components in random models. *Zastos. Mat.* 13, 521–527.

Zusammenfassung

In der Theorie der linearen Modelle kann man die Schätzung von Varianzkomponenten vollständig aus der Schätztheorie für den Mittelwert herleiten; diese 'Streuungs-Mittelwert-Korrespondenz' wird am Beispiel des Hsu-Modells verdeutlicht. Sie ergibt kurze neue Beweise für verschiedene in der Literatur vorkommende Sätze über erwartungstreue invariante quadratische Schätzer mit kleinstem BAYES-Risiko bzw. mit kleinster Varianz. Bei reinen Varianzkomponenten-Modellen wird schließlich erwartungstreue positiv-semidefinite quadratische Schätzbarkeit charakterisiert an Hand der Designmatrizen, die das Modell definieren.

Résumé

Dans ce travail on démontre de nouveau quelques théorèmes concernant l'estimation des composants de la variance dans le modèle linéaire de Hsu: Par l'application d'une 'correspondance dispersion-moyenne' la question posée se trouve réduite aux problèmes déjà connus et résolus dans la théorie d'estimation de la moyenne. Dans les théorèmes cités on étudie des estimateurs quadratiques, invariants, sans biais, qui minimisent soit le risque bayésien soit la variance. En outre, l'existence d'un estimateur quadratique, non-négatif, sans biais d'un seul composant de la variance est caractérisée par des matrices qui, elles, déterminent le modèle.

Received July 1975. Revised May 1976.